ETHNICITY, COMMUNICATION AND GROWTH¤

Jean-Louis ARCAND*,
Patrick GUILLAUMONT** and Sylviane GUILLAUMONT JEANNENEY**

* *CERDI-CNRS, Université d'Auvergne, and CRDE, Université de Montréal*
***CERDI-CNRS, Université d'Auvergne*

In this paper we consider the link often alleged between ethnic diversity and the growth rate of GDP per capita. We first assume that it is ethnic *polarization* rather than ethnic fragmentation that is harmful for growth so that the relationship may be non-linear. Second, we hypothesize that the impact of ethnic diversity on growth may depend on communication costs. This leads us to estimate a traditional growth rate equation on cross sectional data in a switching regression framework. In "low communication costs countries", the relationship between growth and ethnic diversity is U-shaped. On the other hand, in "high communication costs countries", growth is a decreasing function of ethnic diversity and the severity of the latter's deleterious impact is an increasing function of communication costs, proxied here by the illiteracy rate. The regime that a country belongs to is a function of two proxies for communication costs: the illiteracy rate and population density. The impact of ethnic diversity on growth seems not to operate through macroeconomic policy choices. Rather it is a direct transmission mechanism, in which ethnic diversity affects private and public ressource allocation, that appears to dominate.

I. Introduction

Several recent cross-sectional studies in economics have underlined the role of ethnic factors as determinants of the growth performance of developing countries. Mauro (1995), for instance, considers ethnic fragmentation to be a determinant of corruption which in turn may have a negative impact on growth. Collier and Hoeffler (1998) focus on the role of ethnic diversity in triggering civil war. Easterly and Levine (1997), as well as Temple (1998) argue that ethnic fragmentation leads to poor policies that, in turn, affect growth performance. At the outset, it should be noted that several other studies do not confirm the importance of ethnic diversity in cross-country growth regressions (Sachs and Warner, 1997, Rodrik, 1998, Guillaumont, Guillaumont-Jeanneney and Brun, 1998).

The empirical finding that ethnic diversity influences growth raises many questions, not the least of which is what to do about it from the policy perspective. The purpose of this paper is not to contest the fact that ethnicity may play an important role as a determinant of growth. Rather, we wish to investigate *how* ethnicity affects growth. This paper thus considers three issues.

¤ We thank Elliot Berg, Betty Sadoulet, Marielle Deméocq and seminar participants at the CERDI development workshop for useful comments. The usual disclaimer applies.

First, the current literature has been almost exclusively focused on ethnic *diversity*.[1] We argue that, rather than fragmentation *per se*, it is polarization that may constitute a fetter on growth. This is why we use the term "ethnicity" rather than "ethnic fragmentation".

Second, we examine different transmission mechanisms in which ethnicity affects growth. These may be divided into two categories : (i) direct mechanisms, in which ethnicity affects the allocation of ressources by private or public agents, and (ii) indirect mechanisms in which the impact of ethnicity on growth is mediated through its impact on macro policy choices.

Third, we posit that ethnicity does not affect growth in a uniform manner across countries. Rather, its impact on growth depends on the capacity of different social groups to communicate with each other, as might be measured by such indicators as literacy or population density. We test this hypothesis using a two-regime switching regression model of the determinants of growth that distinguishes between countries that are weakly integrated and where communication costs are high, *versus* those where the opposite is true.

II. Fragmentation *versus* Polarization

Given that it is the deleterious impact of ethnic divisions that has been the focus of the recent literature on ethnicity and growth, it is perhaps not surprising that the empirical measure of choice has been an indicator of ethno-linguistic fragmentation, constructed by Taylor and Hudson (1972) on the basis of data published by Soviet geographers at the beginning of the sixties. This indicator (which is by now denoted by most researchers as ELF or $ELF60$) is defined as follows :

$$ELF = 1 - \sum_{i=1}^{i=n} \left(\frac{x_i}{N} \right)^2 ,$$

where x_i is the number of people in the i th ethnic group, $N = \sum_{i=1}^{i=n} x_i$ is the total population and n is the number of ethnolinguistic groups in the country; ELF thus measures the probability that two randomly selected persons from a given country will not belong to the same ethnolinguistic group.[2]

[1] Collier and Hoeffler (1998) and Temple (1998) constitute an exception to this tendency.
[2] Mauro, 1995, p. 692.

The point we wish to raise concerns the distinction between ethnic *diversity* and ethnic *polarization*, two concepts that are often equated but that are in fact quite distinct. Let $\bar{x}$ denote the mean size of the ethnic groups, and let s^2 denote the variance of the size of the ethnic groups. Then it can be readily shown that we may re-express ELF as :

$$ELF(n, s^2, \bar{x}^2) = 1 - \left(\frac{s^2 + \bar{x}^2}{n\bar{x}^2} \right) = 1 - n^{-1}\left(1 + c^2\right),$$

where $c = s / \bar{x}$ is the coefficient of variation. Five comments are in order regarding this measure. First, it is immediate that, for a homogeneous population made up of a single ethnic group, we will have $n = 1$, $s^2 = 0$ and thus $ELF(1,0,\bar{x}^2) = 0$. Second, for a population made up of n ethnic groups of identical size, we will have $s^2 = 0$ and $ELF(n,0,\bar{x}^2) = 1 - n^{-1}$. Third, when one moves from a situation with a single ethnic group to a situation with two ethnic groups of equal size, the measure of ethnic diversity goes from 0 to 0.5. Fourth, consider the two following hypothetical distributions of ethnic groups : one group makes up one half of the population and five other equally-sized groups constitute the rest. The corresponding value of the measure of ethnic diversity is $ELF =$ 0.67. Fifth, the same value is obtained with three ethnic groups of equal size.

It is important to note that the third case, two ethnic groups of equal size, corresponds to what one might term extreme "polarization", and polarization is not necessarily more favorable, in our view, to increasing the growth rate of per capita GDP than is the situation described in the fourth or the fifth case, despite the fact that the indicator of ethnic fragmentation is larger in the fourth or fifth case (0.67) than in the polarized case (0.5).

An operational definition of polarization would be that it corresponds to a situation where the probability of confrontation between ethnic groups is high. This is manifestly the case when there are two ethnic groups of equal size. But it is also the case when there is one dominant group and several smaller groups likely to form a coalition that is capable of opposing the dominant group.

It is intuitively appealing that the maximum level of ethnic polarization is reached when there are two ethnic groups of equal size. In this case, ELF equals one half. It is also intuitively appealing to assume that any measure of ethnic polarization will be smaller than the previously defined maximum for any other configuration of ethnic groups, and thus, at least locally, for any value of ELF that is different from one half. Rather than assuming a particular form for what would be a necessarily arbitrary definition of polarization, we prefer to characterize ethnic polarization in the following manner :

Definition. Any workable definition of ethnic polarization, $P \in [0,1]$, must satisfy the following condition, as a function of ethnic diversity ELF :

$$\text{For } P = P(ELF), ELF \in [0,1], P \in [0,1]: \arg\max_{\{ELF\}} P(ELF) = \frac{1}{2};$$

By the usual argument in terms of a second-order Taylor expansion, such a function may be locally approximated (around $ELF = 0.5$) by a quadratic polynomial. This implies that, if one plots ethnic polarization on the vertical axis *versus* ethnic fragmentation on the horizontal axis, the result is an inverted U-shaped curve that reaches its maximum at $ELF = 0.5$ and displays its minimum value at $ELF = 0$ and $ELF = 1$.

Thus, if it is polarization, rather than diversity *per se*, which constitutes the hindrance to growth, a decrease in the degree of ethnic fragmentation will not necessarily increase the growth rate of per capita GDP. Indeed, if one posits an unambiguously negative relationship between the degree of polarization and per capita GDP growth, the inverted-U-shaped relationship between polarization and ethnic fragmentation (with polarization reaching a maximum for a value of the ethnic fragmentation variable equal to one half) induces a U-shaped relationship between ethnic fragmentation and the growth rate of per capita GDP.

The precise shape of the relationship is not critical here. What is critical is to focus one's attention on the fact that the relationship between ethnic diversity and the growth rate of per capita GDP may not be a linear one, in part because ethnic diversity may be a rather poor proxy for more important factors such as polarization. Indeed, our argument here squares rather nicely with that put forward in a recent paper by Collier and Hoeffler (1998) and by Temple (1998), who argue that polarization corresponds to a situation in which the index of ethnic fragmentation lies in the middle range. Such a configuration, according to Collier and Hoeffler (1998), increases the probability of violent conflict among ethnic groups, namely, the probability of civil war. Our view, however, differs from that of the preceding authors in that we allow for different ethno-economic scenarios as well as for different transmission mechanisms running from ethnicity to growth.

III. Transmission Mechanisms

There are several ways in which ethnicity might affect growth that have been suggested in recent work. We divide the existing arguments into two categories : transmission mechanisms in which ethnic concerns have a direct impact on the growth rate of per capita GDP as far as it renders the allocation of resources less efficient, *versus* indirect mechanisms in which the effect of ethnicity on growth is mediated through some intermediate (usually macroeconomic policy) variables.

Direct Transmission to Resource Allocation

One manner in which ethnicity may affect the growth rate of per capita GDP is through its direct impact (i.e., not mediated through macro policy variables) on the allocation of private or public resources within the economy. There are several mechanisms through which this direct effect may operate.

The private or public allocation of investment projects on the basis of ethnically-derived formulae that are often divorced from rate of return considerations will tend to move the economy away from the efficient frontier. Similarly, discrimination in hiring practices based on ethnic concerns rather than competence may, for instance, result in an inefficient allocation of human capital. Both cases illustrate that ethnicity may lead to market segmentation.

The distortion in resource allocation is exacerbated when there exists ethnically-based specialization of activities. For instance, in Niger, Haousa are devoted to trade and production whereas Djerma make up the civil service. Such a form of specialization induces an inefficient allocation of talent, as well as a plethora of civil servants. In many other countries, small trading activities are fulfilled by foreign minorities (such as Indians in Madagascar and in East Africa) ; the result is that, following political disturbances that involve the minority, markets may be disrupted and trade may become more costly and uncertain.

Another mechanism through which ethnicity impacts the growth rate is linked to the hindrance it imposes on consensus-building related to public expenditures in social sectors. Such lack of consensus may be particularly important when it comes to the provision of those public goods that promote economic growth. The classic example would be the efficiency of the educational system, which is likely to be severely impaired when there is no consensus as to the language of instruction, the content of the curriculum, or the geographical location of facilities. The same might be said of infrastructure construction, which will be a function of the geography of the underlying ethnic groups and will not

necessarily be based upon productivity-enhancing concerns. This effect associated with ethnicity might be labeled as the "public good provision" effect.

As an extreme example of the direct transmission from ethnicity to resource allocation, ethnic diversity can also result in political instability and violence that reduce the level of productive private investment (Collier and Hoeffler, 1998), and the productivity of both public and private investment.

Indirect Transmission through the Impact of Macroeconomic Policies on Incentives

An alternative hypothesis is that ethnic diversity affects the growth rate of GDP per capita through its impact on the pursuit of macroeconomic policies conducive to growth. The mechanism through which the indirect effect of ethnic diversity makes itself felt is that it encourages the adoption of policies associated with rent-seeking activities (Mauro, 1995). Since these rent-seeking activities result in distortions with respect to the first-best optimum, economic growth is thereby deleteriously affected. Examples include foreign currency regulations that lead to the establishment of a parallel market (and thus dual exchange rates), interest rate ceilings that lead to negative real rates of interest, preferential credit policies that target specific sectors or economic agents, restrictions on foreign trade, and the regulation of domestic prices at either the consumer or producer level.

The pursuit of exchange rate unification, currency convertibility, fiscal reform or commercial and financial liberalization must be preceded by the establishment of a policy consensus. And such a consensus may only be reached once there is common knowledge of the concessions that each group is willing to make (Alesina and Drazen, 1991). Indeed, each ethnic group will usually be the beneficiary of a specific form of economic rent, and will fail to internalize the costs that this rent imposes on the other ethnic groups (Shleifer and Vishny, 1993). Policy reform is thus slowed, if not paralyzed, by informational costs (see Schiff, 1995, 1998, for a theoretical model). This informational effect associated with ethnic fragmentation and policy reform might be labeled as the "rent keeping and reform attrition" effect.

Ethnicity and Communication

It is not difficult to see why an ethnically homogeneous population should be conducive to a private allocation of resources that moves one towards the efficient frontier, and to public goods being more efficiently provided. Similarly, a state of ethnic consensus should also lead to a higher likelihood of non-distortionary economic policies being implemented. It is also highly likely that ethnic polarization - the coexistence of two ethnic groups of roughly similar size- will lead to a sharp segmentation of markets as well as to conflicts that will negate attempts at consensus-building. Conversely, when there are a great number of ethnic groups of small size (and the index of ethnic fragmentation is therefore large compared with the polarized case), ethnically-motivated barriers to the efficient allocation of resources may be more difficult to maintain and the risks of conflict may diminish. This first argument justifies our hypothesis, enunciated in the preceding section, that the relationship between ethnic fragmentation and growth may be a U-shaped one.

Our second argument is that the benefits of an increase in ethnic fragmentation (from a starting value strictly greater than 0.5) vanish when informational costs are high since there will be little chance of reconciling heterogeneous preferences regarding the allocation of private resources or the provision of public goods, as well as more resistance to the elimination of economic rents. This would appear to be particularly likely in countries where illiteracy rates are high. In a typical manifestation of this phenomenon, the official language may be spoken and understood in written form by a relatively small proportion of the population. Low population density will exacerbate these tendencies as communication costs will be relatively high, and ethnic groups will often be more isolated geographically. Thus, when combined with high levels of illiteracy and low levels of population density, there will be a tendency for the quadratic term in the relationship between the growth rate of per capita GDP and ethnic fragmentation to vanish.

The preceding two arguments suggest that the relationship between the growth rate of GDP and the level of ethnic fragmentation may depend upon a country's illiteracy rate as well as its population density. Our theoretical construct corresponds to the hypothesis that countries may be divided into two groups, corresponding to a "low communication costs" regime (in which the impact of ethnic diversity on the growth rate is U-shaped), and a "high communication costs" regime (in which the impact of ethnic diversity is unambiguously negative).

Moreover, if our theoretical line of reasoning is valid, it should be the case that the negative effect of ethnic fragmentation on growth in the second "high communication costs" regime is stronger the greater are communication costs, in particular, the greater is the illiteracy rate. This suggests that the

appropriate specification in the "high communication costs" regime involves entering ethnic fragmentation in multiplicative form, multiplied, that is, by the illiteracy rate.

Finally, it is worth noting that this last hypothesis (were it not to be rejected by the data) furnishes an answer to a question posed by Temple (1998) in the conclusion of his paper. Having specified a quadratic relationship between ethnic fragmentation and growth, he finds himself puzzled by the low growth rate of subsaharan African countries, as their level of ethnic fragmentation is particularly high, thus leading his model to predict a higher growth rate for these countries. And this, despite his intuition, as suggested by Easterly and Levine, that the high degree of ethnic fragmentation in Africa constitutes part of the explanation for the low growth rate of these economies.

IV. Empirical Results

Preliminaries

The basic empirical relationship that we estimate corresponds to the by now standard growth regression used by most authors in the literature in which the growth rate of per capita GDP is related to (i) the initial level of per capita GDP (expressed in logarithmic terms), (ii) the level of human capital as measured by the average number of years of schooling (we use the measure constructed by Barro and Lee (1993)), (iii) a decade dummy (we confine our attention to the 1970s and 1980s). The fourth explanatory variable that we introduce is the index of ethnic fragmentation, which we introduce in linear (column 1), as well as in quadratic (column 2) form. When ethnicity enters in purely linear form, we are in a situtation in which the impact of ethnicity on growth is unambiguously negative. On the other hand, when ethnicity enters in quadratic form, we are in a case in which the impact of ethnicity on growth is based on polarization concerns. The first case is that considered by Easterly and Levine (1997), whereas the second is close in spirit to the specification implemented by Temple (1998). Observations are pooled by countries over two decades and standard errors were computed using White's heteroskedasticity-consistent method.

A first preliminary test of our hypothesis based on communication costs is given by the results of two Chow tests that are presented in the lower part of the first two columns of Table I. In the first

Chow test, we construct a subsample of observations corresponding to those countries with an illiteracy rate in the fourth quartile *and* with a population density in the first quartile of the full sample. Given our theoretical arguments, these observations should correspond to countries in which communication costs are particularly high. The second Chow test is based upon a subsample in which we consider observations corresponding to the fourth quartile of the *ratio* of the illiteracy rate to the population density. For both these sample splits, the p-values of the corresponding tests are extremely small (0.054 and 0.007, respectively, for the linear specification —results are similar for the quadratic specification presented in column 2), indicating that one can readily reject the null-hypothesis of parameter constancy across the two subsamples.

A second test of our ethnicity and communication cost hypothesis is given by three specifications where either the illiteracy rate or the logarithm of the population density (or both variables) appear in multiplicative form. If we focus our attention on column 5 in which the impact of ethnolinguistic fragmentation on growth is allowed to vary as a function of both the illiteracy rate and the logarithm of population density, it is obvious, given the p-value on the joint significance test (0.042), that it would be erroneous to assume that the impact of ethnic fragmentation on growth is independent of communication costs.

While the results of the Chow tests as well as the multiplicative specifications do suggest that the relationship between the growth rate of per capita GDP and ethnic fragmentation is a function of our communication variables, a more sophisticated approach is needed in order to directly test our hypothesis.

A Two Regime Specification

In order to condition the relationship between ethnic diversity and the growth rate of GDP per capita in a manner that corresponds to the theoretical arguments enunciated above, we turn to a switching regression specification in which regime 1 corresponds to the "low communication costs" scenario and regime 2 corresponds to the "high communication costs" scenario. Whether a country belongs to regime 1 or regime 2 will depend simultaneously upon its illiteracy rate and its population density.

Let x_i ($i = 1,...,N$, indexes observations) denote the $N \times 4$ matrix of control variables that is common to both regimes (a constant term, the initial level of GDP per capita, a decade dummy, and the Barro-Lee measure of human capital). Let $x_{1i} = [x_i \;\; ELF60_i \;\; ELF60_i^2]$ denote the $N \times 6$ matrix of explanatory variables that correspond to regime 1 and $x_{2i} = [x_i \;\; ELF60_i \times ILLITERACY_i]$ denote the $N \times 5$ matrix of explanatory variables that corresponds to regime 2; the dependent variable (the growth rate of per capita GDP) will be denoted by y_i with y_{1i} denoting the growth rate under regime 1 and y_{2i} denoting the growth rate under regime 2. The econometric specification is then given by the following system of equations:

$$\text{Regime } 1: y_{1i} = b_1' x_{1i} + u_{1i}$$
$$\text{Regime } 2: y_{2i} = b_2' x_{2i} + u_{2i}$$

where the choice between regimes is given by the following sorting condition

$$y_i = \begin{cases} y_{1i} & if \;\; g'z_i + v_i < 0 \\ y_{2i} & if \;\; g'z_i + v_i \geq 0 \end{cases}$$

with $v_i \sim N(0,1)$, and where z_i is an $N \times 3$ matrix constituted by a constant term, the illiteracy rate, and the logarithm of population density. The distributional assumptions on the disturbance terms in the two regimes are given by $(u_{1i}, u_{2i}) \sim N(0,0,\Sigma)$. That is, the disturbance terms in the two equations are distributed according to the bivariate normal density with zero means and covariance matrix given by

$$\Sigma = \begin{bmatrix} s_1^2 & s_{12} \\ s_{12} & s_2^2 \end{bmatrix}.$$

It is well known (e.g., Maddala (1983), Quandt (1988)) in the context of this type of model that the off-diagonal term s_{12} is not identifiable. Note that, for identification purposes (Maddala and Nelson (1975), p. 424), we must (i) normalize the standard deviation of the selection equation such that it is equal to one (that is why v_i is distributed $N(0,1)$ and not $N(m_v, s_v^2)$) and (ii) normalize the

coefficient on the constant term in the selection equation to one (i.e., the parameters in the selection equation are only identified up to a multiplicative constant).

The selection equation which determines whether a country belongs to regime 1 or regime 2 is given by a (latent) regime indicator function defined as follows :

$$I_i(z_i) = \begin{cases} 1 \; if \;\; g'z_i + v_i < 0 \\ 0 \;\; if \;\; g'z_i + v_i \geq 0 \end{cases}$$

We can therefore write our left-hand-side variable from the growth equations in the following form:

$$y_i = I_i(z_i)b_1'x_{1i} + (1 - I_i(z_i))b_2'x_{2i} + I_i(z_i)u_{1i} + (1 - I_i(z_i))u_{2i},$$

where we approximate $I_i(z_i)$ in continuous form by the probit function :

$$\hat{I}_i(z_i) = \int_{-\infty}^{g'z_i} \left(\frac{\exp\{-v^2/2\}}{\sqrt{2p}} \right) dv.$$

One can then easily construct the corresponding likelihood function and maximize it with respect to b_1, b_2, s_1, s_2, g which, owing to the above normalizations, are all identifiable. Note that this procedure does not arbitrarily assign an observation to a given regime : this process is carried out optimally through the maximization of the likelihood function, thus allowing the data (conditioned by the variables included in the selection equation) to sort themselves freely into the two regimes.

The results are presented in Table II. In the first column, the difference between the two regimes is that ethnic diversity enters in quadratic form in regime 1 and in linear form in regime 2. In column (2), ethnic diversity enters in multiplicative form, as our formal hypothesis would have it, in regime 2. If we begin by considering the selection equation, it is worth pointing out that the coefficients on the illiteracy rate and on population density have the signs predicted by our theory (a positive sign indicates a higher probability of the observation falling under regime 2). Moreover, in column (2) ethnic diversity has the appropriate U-shape under regime 1, and both coefficients are statistically significant at the usual confidence levels, while ethnic diversity multiplied by the illiteracy rate in regime 2 is negative and statistically significant.

These results offer strong support for our main theoretical hypothesis that the impact of ethnic diversity on the growth rate of per capita GDP is conditioned by communication costs. Remarkably, for this type of switching regression, sample separation into the two regimes is crisp (with highly significant coefficients on our proxies for communication costs in the selection equation), and despite the difference between the two regimes being minimal (from the econometric point of view), the coefficients on ethnic fragmentation and ethnic fragmentation squared in regime 1, and on ethnic fragmentation times the illiteracy rate in regime 2 are estimated rather precisely.

Figure 1 illustrates the results presented in column (2) of Table II. The U-shaped curve represents the predicted impact of ethnic fragmentation on the growth rate of per capita GDP for those observations belonging to regime 1, once the growth rate has been purged of the effects of the usual control variables. The straight lines represent the predicted value of the growth rate of GDP per capita for observations belonging to regime 2, for different values of the illiteracy rate. The middle straight line corresponds to the median value of the illiteracy rate of countries that belong to regime 2. The uppermost and lowermost straight lines correspond to a illiteracy rates equal to the cutoff value between the first and second quartiles, and the third and fourth quartiles, respectively. Note that the minimum value of the predicted growth rate of per capita GDP for countries that belong to regime 1 is attained at ELF60 near 0.4 —not far from our theoretically motivated value of 0.5.

From the empirical results in Table II we can construct the predicted probability of a given observation (country) belonging to regime 1 or regime 2. Since the true indicator function is approximated by the probit function, these probabilities will not appear in discrete (i.e., 0 or 1) form, that is, they will often be strictly comprised between zero and one. Table III presents the probability of various groupings of countries belonging to regime 2.

Macroeconomic Policy Factors in the High Communication Costs Regime: Direct or Indirect Transmission?

In order to study whether the impact of ethnic diversity on growth operates through the direct or indirect (i.e., policy-mediated) mechanisms posited above, we next regressed several standard macroeconomic policy variables commonly used in the empirical growth literature, on ethnic diversity. We did so in two different ways, first, for those observations identified by the switching growth regression as belonging to the "high communication costs" regime (i.e., regime 2 which accounts for

77% of observations) and, second, using a switching regression procedure similar to that used for the growth regressions.

Our choice of macroeconomic policy variables, the black market premium, the fiscal surplus and financial depth, was in large measure determined by previous work in this area (see Easterly and Levine, 1997, as well as Sachs and Warner, 1996).[3] The black market premium is a good indicator of those distortions induced by foreign exchange restrictions and the misallignment of exchange rates. A fiscal deficit can be interpreted as a symptom of an inability to implement adequate policies and is often associated with significant distortions stemming from tariff and taxation policies. Low financial depth, measured by the ratio of liquid assets to GDP, is often taken to be the result of financial repression, but can also reflect individual saving behavior independent of macroeconomic policies (for instance, a reluctance to make deposits in banks controlled by other ethnic groups and a preference for informal and ethno-specific credit associations).

First, consider the results based on the subsample of those countries classified by our switching growth regression procedure (column (2) of Table II) as belonging to regime 2. In essence, we thus carried out a Heckman procedure where we control for selection bias by introducing, as an additional explanatory variable, the predicted probability of belonging to regime 2 computed in column (2) of Table 2. The empirical results corresponding to this procedure are reported in Table IV.[4] It is apparent from the results that it is financial depth that is significantly and negatively affected by ethnic diversity for countries belonging to regime 2 (note that the coefficient on the predicted probability of the observation belonging to regime 2 is also statistically significant at the usual levels of confidence), while the impact of ethnic diversity on the fiscal surplus and on the premium on the black market exchange rate premium is statistically insignificant. If one interprets financial depth as a "structural", rather than as a "policy" variable, as argued above, then it is more indicative of a direct rather than an indirect effect of ethnic diversity on the growth rate.

[3] The sample used in these estimations is necessarily smaller than that used in our basic switching regression specification in that there are many observations for which the policy variables in question are missing.

[4] As we were seeking to maximize sample size, the relevant sample here includes a number of observations which were not included in the sample in column (2) of Table II, but for which we could nevertheless (i) construct the estimated probability of being in regime 2 and (ii) perform the regression of the policy variables on ethnic diversity while controlling for selection bias. The results do not differ appreciably if we confine our attention to those observations corresponding to regime 2 in the switching regression results reported in column (2) of Table II and for which the policy variables are also available.

A more formal empirical argument can be used to reject the hypothesis that it is the indirect effect that provides the mechanism through which ethnic diversity affects the growth rate of GDP per capita, by investigating whether the impact of ethnic diversity on the policy variables differs by regime within a switching regression procedure similar to that used for the growth regressions. In regime 2, as in the growth regression presented in column (2) of Table II, ethnic diversity enters in multiplicative form (multiplied by the illiteracy rate, that is) while it enters in quadratic form under regime 1. The selection equation is of the same form as above and includes the illiteracy rate and the logarithm of population density as explanatory variables. Of related interest is whether the separation of countries into the two regimes within the switching policy regression framework yields the same sample separation as in the case of the switching growth regressions. In this case one might be tempted to conclude that the differential impact of ethnic diversity on growth operates through its differential impact on the policy variables which, in turn, determine the growth rate of GDP per capita. One should then interpret the results presented in the first two columns of Table II as being reduced form equations that correspond to a more complex structural specification.

Our results for the fiscal surplus (column (5) of Table II) may be summarized as follows. The average probability of an observation falling under regime 2 in the fiscal surplus equation is equal to 0.427 (the corresponding figure for the growth regression presented in column (2) of the same Table is 0.771), the illiteracy rate is insignificant in the corresponding selection equation, and the three coefficients on ethnic fragmentation in the structural equations are statistically insignificant. The correlation between the probability of being in regime 2 in the switching growth system (column (2) of Table II) and of being in regime 2 in the switching fiscal surplus system (column (5) of Table II) is equal to 0.502. The sample separation is therefore manifestly not the same, by any stretch of the imagination.

In the black market premium equation, on the other hand, the sample separations are closer (see column (4) of Table II). The coefficients on the illiteracy rate and population density in the selection equation are of the correct sign and statistically significant (the illiteracy rate is marginally so). The correlation between the probability of an observation being in regime 2 in the growth system and being in regime 2 in the black market premium system is equal to 0.515. The average probability of being in regime 2 is equal to 0.525. However, none of the coefficients on ethnic diversity are statistically significant.

In contrast, financial depth (column (3) of Table II) gives rather different results, at least in terms of the estimated policy equation for regime 2, in which ethnic diversity times the illiteracy rate is

statistically significant (t = -2.69). The average probability of an observation falling under regime 2 is equal to 0.548, and the correlation between the probability of an observation being in regime 2 in the growth system and of the same observation being under regime 2 in the policy system is equal to 0.685. While these results are reassuring in the sense that they confirm that the separation of the sample into "low communication costs" and "high communication costs" countries is not entirely spurious, the coefficients on the quadratic expression in ethnic diversity under regime 1 are statistically insignificant.

It would appear, therefore, that the differential impact of ethnic diversity on the growth rate of GDP per capita in the two regimes does not obtain through a differential impact of ethnic diversity on macroeconomic policy choices that, in turn, affect the incentives facing economic agents. Rather, there is a direct effect of ethnic diversity on the growth rate of GDP that differs according to whether a country is in the "low communication costs" versus the "high communication costs" regime.

It should be noted that these results are in contrast with Easterly and Levine (1997) and Temple (1998) who both assume that ethnic diversity affects growth through its impact on policy variables. Our paper generalizes their results and combines them, in the sense that Easterly and Levine introduce ethnic fragmentation in linear form, whereas Temple introduces it in quadratic form. Neither Easterly and Levine nor Temple, however, allow for a differential impact of ethnic diversity based on a third group of variables.

V. Concluding Remarks

The contribution of this paper about the impact of ethnic diversity on growth lies in (i) our having gone beyond the usual linear specification adopted by most authors, and (ii) our having shown that this impact differs for different categories of countries. We first assume that it is ethnic *polarization* rather than diversity which is harmful for growth, so that the relationship between growth and ethnic diversity may be non-linear. We also assume that the impact of ethnic diversity depends on the level of communication costs within a country, which leads to differenciate countries according to this level.

We investigated these hypotheses in two successive steps. First, using simple Chow tests and multiplicative specifications, we considered whether the regression coefficients in the basic specifications of the determinants of growth were (i) stable across subsamples constructed on the basis of differing levels of communication costs, (ii) a function (in the case of the coefficient on ethnic fragmentation) of

communication costs. We proxied these communication costs by the illiteracy rate and the population density.

Second, not having rejected the null hypothesis that the cost of communication is a significant determinant of the impact of ethnic diversity on growth, we considered a switching regression model with exogenous switching, in which regime 1 corresponds to "low communication costs countries" while regime 2 corresponds to "high communication costs countries". In regime 1, ethnic diversity enters in quadratic form, and while the expected sign on ethnic diversity is negative, the expected sign on ethnic diversity, squared, is positive, implying a U-shaped relationship between the growth rate of GDP per capita and ethnic diversity, conditional on the usual control variables. In regime 2, on the other hand, ethnic diversity enters in multiplicative form (multiplied by the illiteracy rate), and the expected sign of the coefficient is negative. Our selection equation, for its part, is a function of our two proxies for communication costs.

Finally, in the concluding section of part IV, we investigated the differential impact of ethnic diversity on three policy variables commonly used in the empirical growth literature. We did so, first, by carrying out what essentially boils down to a Heckman procedure : for a subsample of observations classified by our switching growth regression as belonging to regime 2 (the high communication costs regime), we regressed each policy variable on ethnic diversity while including the predicted probability (computed from the selection equation of the switching growth regression) of the observation belonging to that regime. Second, we estimated a switching regression model in which the dependent variable is the policy variable, and where the selection equation is similar to the one used in the growth regressions. The regimes were distinguished by ethnic diversity entering in linear (multiplicative) versus quadratic form. Our results, which showed (i) that the policy variables were statistically unrelated to ethnic diversity when one controls for sample selection bias, and (ii) that the sample separation given by the policy regressions is far from that given by the growth regressions, imply that one cannot attribute the impact of ethnic diversity on growth to an indirect effect that operates through policy choices. Rather, it is the direct transmission mechanism that dominates.

The policy implications of our results, in contrast to other recent papers (e.g. Easterly and Levine, 1997), are not so depressing. When the impact of ethnic diversity on growth is unambiguously negative, a high of degree of ethnic diversity would appear to be a question of fate, in the sense that it inexorably lowers the growth rate of GDP per capita. In our results, on the other hand, ethnic diversity, for countries that belong to regime 2, is a severe handicap that can, however, be overcome. Indeed,

through literacy programs and improvements in infrastructure possibly supported by external assistance, a country should be capable, first (while remaining within regime 2), of dampening the deleterious effects of ethnic fragmentation (reducing the absolute value of the slope of the straight line in Figure 1) and, later, of switching over to what we have deemed regime 1, the "low communication costs regime". Having operated this switch, ethnic diversity, rather than constituting a fetter on growth, can become an asset.

Ethnicity, we have shown, is not important because of its effect through macroeconomic policies. Therefore, even with a high degree of ethnic diversity, there is nothing that condemns a country to pursuing poor policies. Of course, there may be other aspects of policy, that are not captured by the three proxies commonly used in the empirical growth literature (black market premium, fiscal surplus and financial depth), and that are affected by ethnic diversity.

References

ALESINA, A. AND DRAZEN, A. (1991), "Why are Stabilizations Delayed," *American Economic Review* 81:1170-1188.

ATLAS NARODOV MIRA (1964) (Moscow : Miklukho-Maklai Ethnological Institute at the Department of Geodesy and Cartography of the State Geological Committee of the Soviet Union, 1964).

COLLIER, P., AND HOEFFLER, A. (1998), "On Economic Causes of Civil War," processed, Center for the Study of African Economies, Oxford University, January.

EASTERLY, W., AND LEVINE, R. (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, November, 1203-1250.

GUILLAUMONT, P., GUILLAUMONT-JEANNENEY, S., AND J.- F. Brun (1998), "How Instability Lowers African Growth," *Journal of African Economies*, forthcoming.

MADDALA, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics* (Cambridge : Cambridge University Press).

MADDALA, G. S. AND NELSON, F. D. (1975), "Switching Regression Models with Exogenous and Endogenous Switching," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 423-426.

MAURO, P. (1995), "Corruption and Growth," *Quarterly Journal of Economics* 110:681-712.

QUANDT, R. (1988). *The Econometrics of Disequilibrium* (Oxford : Basil Blackwell).

RODRIK, D. (1998), "Trade Policy and Economic Performance in Sub-Saharan Africa," NBER Working Paper No. 6562.

SACHS, J., AND WARNER, A. (1997), "Sources of Slow Growth in African Economies," *Journal of African Economies* 6:335-376.

SCHIFF, M. (1995), "Capital socioculturel, diversité ethnique et réforme économique en Afrique subsaharienne," *Revue d'Économie du Développement* 3(3):41-53.

SCHIFF, M. (1998), "Ethnic Diversity and Economic Reform in Sub-Sahara Africa," *Journal of African Economies* 7(3):348-362.

SHLEIFER, A., AND VISHNY, R. (1993), "Corruption", *Quarterly Journal of Economics* 108:599-617.

TAYLOR, C. , AND HUDSON, M. (1972), *World Handbook of Political and Social Indicators*, 2nd edition (New Haven : Yale University Press).

TEMPLE, J. (1998), "Initial Conditions, Social Capital and Growth in Africa," *Journal of African Economies* 7(3):309-347.

TABLE I

BASELINE SPECIFICATION AND TESTS OF ROBUSTNESS

(DEPENDENT VARIABLE : GROWTH RATE OF GDP PER CAPITA)

	(1)	(2)	(3)	(4)	(5)
Constant	0.007	0.015	0.025	0.005	0.019
	(0.31)	(0.62)	(1.02)	(0.23)	(0.75)
Dummy for the 70s	0.019	0.019	0.018	0.019	0.019
	(5.33)	(5.32)	(5.25)	(5.52)	(5.43)
Log of initial income	-0.0004	-0.002	-0.001	0.0001	-0.0009
	(-0.11)	(-0.56)	(-0.47)	(0.04)	(-0.24)
Log of schooling	0.007	-0.008	0.002	0.005	0.001
	(1.31)	(-1.45)	(0.33)	(0.90)	(0.24)
Ethnic fragmentation	-0.018	0.020	-0.001	-0.036	-0.020
	(-2.46)	(0.74)	(-0.09)	(-2.67)	(-1.15)
Ethnic fragmentation, squared		-0.047			
		(-1.50)			
Multiplicative variables					
Ethnic fragmentation			-0.034		-0.026
× illiteracy rate			(-2.17)		(-1.63)
Ethnic fragmentation				0.005	0.004
× log of population density				(1.45)	(1.19)
Test on joint significance of multiplicative terms (p-value)					$\chi^2(2)=6.331$ (0.042)
Chow tests					
Chow test†: subsample of countries belonging to first quartile of population density *and* fourth quartile of illiteracy (p-value)	$\chi^2(5)=10.83$ (0.054)	$\chi^2(6)=9.03$ (0.171)			
Chow test†: subsample of countries belonging to fourth quartile of the ratio of illiteracy on population density (p-value)	$\chi^2(5)=15.66$ (0.007)	$\chi^2(6)=13.01$ (0.042)			
R-squared	0.224	0.237	0.241	0.246	0.225
No. of observations	157	157	157	157	157

Note: White heteroskedasticity-consistent t-ratios in parentheses, unless otherwise noted.

Datasource : growth rate of GDP per capita, log of initial income, log of schooling, and ethnic fragmentation, Easterly and Levine (1997); Illiteracy rate and population density : World Bank data.

TABLE II

SWITCHING REGRESSION WITH EXOGENOUS SWITCHING AND UNKNOWN SAMPLE SEPARATION (T-STATISTICS IN PARENTHESES)

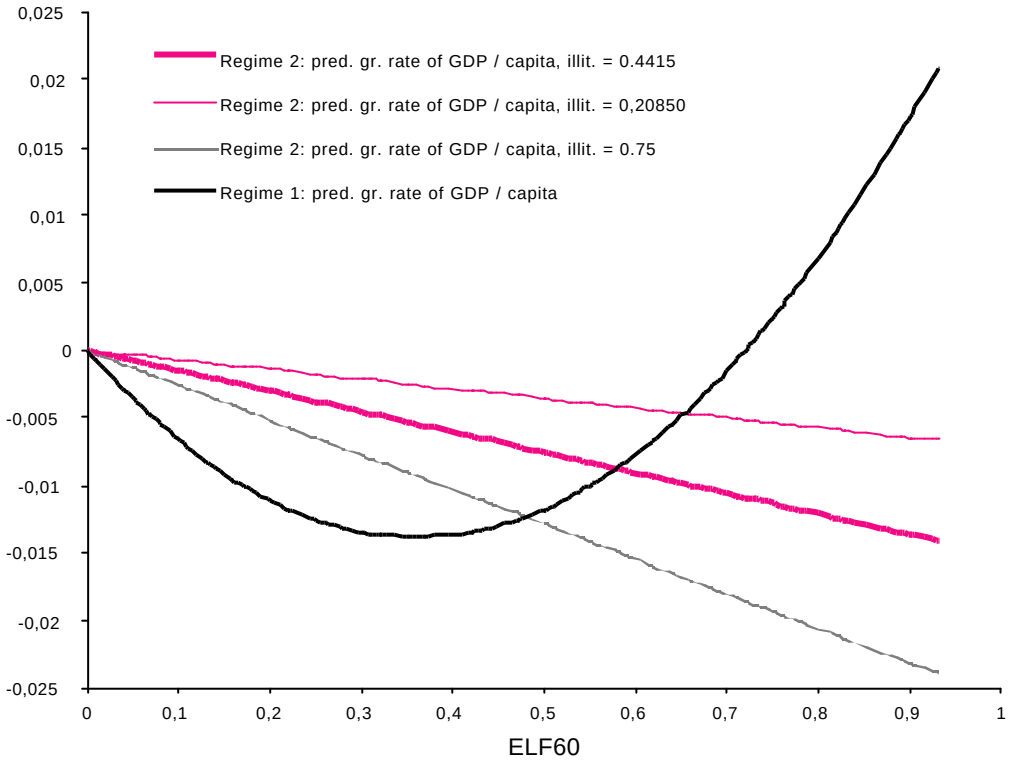
			Policy variable regressions		
	1	2	3	4	5
Selection equation					
illiteracy rate	24.81	27.21	2.57	0.504	984.9
	(2.22)	(2.12)	(4.56)	(1.70)	(0.0003)
population density	-0.735	-0.778	-0.474	-0.306	-102.0
	(-3.10)	(-3.05)	(-4.59)	(-7.57)	(-0.0009)
Regime 1					
Dependent variable	Gr. rate of GDP / capita	Gr. rate of GDP / capita	Financial depth	Black mkt. Premium	Fiscal surp./ GDP
constant	0.200	0.205	0.638	0.022	-0.045
	(7.08)	(7.32)	(8.28)	(2.01)	(-4.28)
elf60	-0.077	-0.077	-0.129	0.0001	-0.039
	(-2.89)	(-2.84)	(-0.26)	(0.001)	(-0.42)
elf60, squared	0.105	0.107	-0.038	0.004	0.114
	(2.31)	(2.29)	(-0.04)	(0.06)	(0.76)
dummy for 1970s	0.003	0.003			
	(1.03)	(0.99)			
log of schooling	0.005	0.005			
	(0.68)	(0.63)			
log GDP/capita	-0.019	-0.020			
	(-5.55)	(-5.54)			
σ_1	0.006	0.006	0.247	0.031	0.040
	(6.13)	(6.58)	(10.03)	(5.47)	(10.67)
Regime 2					
Dependent variable	Gr. rate of GDP / capita	Gr. rate of GDP / capita	Financial depth	Black mkt. Premium	Fiscal surp./ GDP
constant	-0.001	0.020	0.271	0.349	-0.043
	(-0.02)	(0.53)	(13.41)	(2.25)	(-4.36)
elf60	-0.016				
	(-1.54)				
elf60 × illiteracy rate		-0.034	-0.125	0.373	-0.019
		(-2.18)	(-2.69)	(1.17)	(-0.704)
dummy for 1970s	0.022	0.022			
	(4.66)	(4.54)			
log of schooling	0.004	-0.001			
	(0.58)	(-0.16)			
log GDP/capita	0.0006	-0.0009			
	(0.12)	(-0.18)			
σ_2	0.024	0.024	0.097	0.428	0.042
	(17.50)	(18.04)	(7.61)	(12.07)	(17.93)
average prob. of regime 2	0.770	0.771	0.548	0.525	0.427
correlation of prob. of regime 2 with prob. of regime 2 in column (2)	n.a.	n.a.	0.685	0.515	0.502
log of likelihood function	399.33	401.29	39.15	39.56	230.60
number of observations	157	157	154	157	131

Note : Maximum likelihood estimation.

FIGURE 1

SWITCHING REGRESSION WITH UNOBSERVABLE SAMPLE SEPARATION
LINEAR-MULTIPLICATIVE (WEAKLY INTEGRATED) AND QUADRATIC (HIGHLY INTEGRATED) REGIMES

(Growth rate of GDP per capita conditional on constant, decade dummy, linear convergence effect and schooling on vertical axis)



Note : U-shaped curve corresponds to predicted value of growth rate as a function of ELF60 in highly integrated regime (regime 1). Middle straight line corresponds to predicted value of growth rate in weakly integrated regime (regime 2) estimated at the median value of illiteracy for observations belonging to regime 2. Upper and lower straights lines represent predicted value of growth rate estimated for illiteracy rate equal to the limit value separating the first from the second quartile, and the third from the fourth quartile, respectively.

TABLE III

PREDICTED PROBABILITIES OF BELONGING TO REGIME 2, BY GROUPING

Subsample	No. Obs.	Mean estimated probability of being in regime 2	Standard deviation
Full sample	157	0.771	0.373
African	43	0.993	0.045
Non-African	114	0.687	0.407
Latin American	42	0.880	0.255
1st quartile of GDP per capita	39	1.000	0.000
4th quartile of GDP per capita	39	0.227	0.299

Note : computed from the estimated value of the probability of belonging to regime 2 based on the switching regression presented in column (2) of Table II.

TABLE IV

CORRELATION BETWEEN POLICY VARIABLES AND ETHNIC DIVERSITY
UNDER THE "HIGH COMMUNICATION COSTS / WEAKLY INTEGRATED" REGIME

Dependent variable	Financial depth	Black market premium	Fiscal surplus / GDP
Constant	0.920	-0.347	-0.092
	(4.82)	(-1.62)	(-1.45)
Ethnic fragmentation	-0.158	0.058	-0.008
	(-2.30)	(0.45)	(-0.57)
Predicted probability of being in Regime 2	-0.536	0.633	0.049
	(-2.61)	(2.53)	(0.76)
R-squared	0.126	0.023	0.011
No. of observations	119	122	105

Note : White heteroskedasticity-consistent t-ratios in parentheses. Predicted probability of being in regime 2 computed from column (2) of Table II.